

## COMPOSITIONS UTILISABLES POUR REGULER L'ACTIVITE

### DE LA PARKINE

La présente invention concerne des compositions et méthodes utilisables pour réguler l'activité de la parkine. Elle concerne notamment une nouvelle protéine, désignée PAP1, partenaire de la parkine, ainsi que les peptides ou polypeptides dérivés ou homologues de cette protéine. Elle concerne également des composés capables de moduler au moins partiellement l'activité de la Parkine, notamment d'interférer avec l'interaction entre la parkine et PAP1. La présente invention est utilisable dans les domaines thérapeutiques, diagnostiques ou pour la constitution de cibles pharmacologiques permettant le développement de nouveaux médicaments.

Le gène de la Parkine est muté dans certaines formes familiales (juvéniles autosomiques récessives) de la maladie de Parkinson (Kitada *et al.*, 1998). La maladie de Parkinson (Lewy, 1912) est une des maladies neurodégénératives les plus communes, affectant plus de 1% de la population de plus de 55 ans. Les patients atteints de cette maladie ont des troubles neurologiques rassemblés sous le terme de syndrome parkinsonien, caractérisé par une rigidité, une bradykinésie, et un tremblement de repos. Ces symptômes sont la conséquence d'une dégénérescence des neurones dopaminergiques de la substance noire du cerveau.

La plupart des cas avec une maladie de Parkinson n'ont pas d'histoire familiale. Cependant, il existe des cas familiaux dont certains correspondent à une forme monogénique de la maladie. A l'heure actuelle, seulement trois gènes différents ont été identifiés dans certaines formes héréditaires rares. La première forme correspond à une forme autosomique dominante dont le gène responsable code pour l'alpha Synucléine (Polymeropoulos *et al.*, 1997). Cette protéine est un constituant abondant des inclusions intracytoplasmiques, appelées corps de Lewy, qui servent de marqueur de la maladie de Parkinson (Lewy, 1912). La seconde forme, également autosomique dominante, est liée à une mutation dans un gène codant pour une hydrolase appelée ubiquitin carboxy-terminal hydrolase L1 (Leroy *et al.*, 1998). Cette

enzyme est supposée hydrolyser les polymères ou les conjugués d'ubiquitines en monomères d'ubiquitine. La troisième forme se distingue des précédentes par une transmission autosomique récessive et un début souvent avant 40 ans ainsi que par une absence de corps de Lewy. Ces malades répondent plus favorablement à la levodopa, un précurseur de la dopamine utilisé comme traitement de la maladie de Parkinson. Le gène impliqué dans cette forme code pour une nouvelle protéine appelée Parkine (Kitada *et al.*, 1998).

Le gène de la Parkine est constitué de 12 exons couvrant une région génomique de plus de 500 000 paires de bases sur le chromosome 6 (6q25.2-q27). A l'heure actuelle, deux grands types de mutation de ce gène, à l'origine de la maladie, sont connus, soit des délétions de taille variable dans la région qui couvre les exons 2 à 9, soit des mutations ponctuelles qui produisent l'apparition prématurée d'un codon stop ou le changement d'un acide aminé (Kitada *et al.*, 1998; Abbas *et al.*, 1999; Lucking *et al.*, 1998; Haltori *et al.*, 1998). La nature de ces mutations et le mode de transmission autosomique récessif suggèrent une perte de fonction de la Parkine conduisant à la maladie de Parkinson.

Ce gène est exprimé dans un grand nombre de tissus et notamment dans la substance noire. Il existe plusieurs transcrits correspondant à ce gène qui proviennent d'épissages alternatifs différents (Kitada *et al.*, 1998; Sunada *et al.*, 1998). Dans le cerveau, on trouve deux types d'ARN messagers dont un est dépourvu de la partie correspondant à l'exon 5. Dans les leucocytes, des ARN messagers de la Parkine ne contenant pas la région codant pour les exons 3, 4 et 5 ont été identifiés. Le plus long des ARN messagers de la Parkine, présent dans le cerveau, contient 2960 bases et code pour une protéine de 465 acides aminés.

Cette protéine a une faible homologie dans sa partie N terminale avec l'ubiquitine. Sa moitié C-terminale contient deux motifs « ring finger » séparés par un domaine IBR (In-Between Ring) correspondant à une région riche en cystéine pouvant

lier des métaux comme les domaines « zing finger » (Moretti, 1999). Par immunocytochimie, il a été montré que la Parkine est localisée dans le cytoplasme et l'appareil de Golgi de neurones de la substance noire qui contiennent de la mélanine (Shimura *et al.*, 1999). De plus, cette protéine est présente dans certains corps de  
 5 Lewy de Parkinsoniens. La fonction cellulaire de la Parkine n'est pas encore démontrée, mais elle pourrait jouer un rôle de transporteur dans les vésicules synaptiques, dans la maturation ou dégradation des protéines et dans le contrôle de la croissance, de la différenciation ou du développement cellulaire. Dans les formes autosomiques récessives juvéniles, la Parkine est absente, confirmant ainsi que la  
 10 perte de cette fonction est responsable de la maladie.

L'élucidation du rôle exact de la protéine Parkine dans le processus de dégénérescence des neurones dopaminergiques constitue donc un enjeu majeur pour la compréhension et l'approche thérapeutique de la maladie de Parkinson et plus  
 15 généralement des maladies du système nerveux central.

La présente invention réside dans l'identification d'un partenaire de la parkine, interagissant avec cette protéine dans des conditions physiologiques. Ce partenaire représente une nouvelle cible pharmacologique pour la fabrication ou la recherche de  
 20 composés capables de moduler l'activité de la parkine, notamment son activité sur la dégénérescence des neurones dopaminergiques et/ou le développement de pathologies nerveuses. Cette protéine, les anticorps, les acides nucléiques correspondant ainsi que les sondes ou amorces spécifiques, sont également utilisables pour la détection ou le dosage des protéines dans des échantillons biologiques, notamment des échantillons  
 25 de tissu nerveux. Ces protéines ou acides nucléiques sont également utilisables dans des approches thérapeutiques, pour moduler l'activité de la parkine, ainsi que tout composé selon l'invention, capables de moduler l'interaction entre la parkine et les polypeptides de l'invention.

La présente invention résulte plus particulièrement de la mise en évidence par  
 30 la demanderesse d'une nouvelle protéine humaine, désignée PAPI (Parkine

Associated Protein 1) ou LY111, interagissant avec la parkine. La protéine PAPI (séquence SEQ ID NO:1 ou 2) présente une certaine homologie avec les synaptotagmines et est capable d'interagir plus particulièrement avec la région centrale de la parkine (représentée sur la séquence SEQ ID NO:3 ou 4). La protéine PAPI a également été clonée, séquencée et caractérisée à partir de différents tissus d'origine humaine, notamment de poumon (SEQ ID NO: 12, 13) et de cerveau (SEQ ID NO: 42, 43), ainsi que des formes courtes, correspondant à des variants d'épissage (SEQ ID NO: 14, 15, 44, 45).

La présente invention résulte également de l'identification et de la caractérisation de régions particulières de la protéine PAPI, impliquées dans la modulation de la fonction de la parkine. La mise en évidence de l'existence de cette protéine et de régions impliquées dans sa fonction permet notamment de préparer de nouveaux composés et/ou compositions utilisables comme agents pharmaceutiques et de développer des méthodes industrielles de criblage de tels composés.

Un premier objet de l'invention concerne donc des composés capables de moduler, au moins partiellement, l'interaction entre la protéine PAPI (ou ses homologues) et la parkine (notamment la parkine humaine), ou d'interférer au niveau de l'interaction entre ces protéines.

Un autre objet de l'invention réside dans la protéine PAPI, ses fragments, dérivés et homologues.

Un autre aspect de l'invention réside dans un acide nucléique codant la protéine PAPI, ses fragments, dérivés ou homologues, ainsi que tout vecteur comprenant un tel acide nucléique, toute cellule recombinante contenant un tel acide nucléique ou vecteur et tout mammifère non-humain comprenant dans ses cellules un tel acide nucléique.

L'invention concerne aussi des anticorps capables de lier la protéine PAPI, ses fragments, dérivés et homologues, notamment des anticorps polyclonaux ou monoclonaux, plus préférentiellement des anticorps capables de lier la protéine PAPI et d'inhiber au moins partiellement son interaction avec la parkine.

Un autre aspect de l'invention concerne des sondes ou amorces nucléotidiques, spécifiques de PAP1, utilisables pour détecter ou amplifier le gène de pap1 ou une région de celui-ci dans tout échantillon biologique.

L'invention concerne encore des compositions pharmaceutiques, des méthodes de détection d'anomalies génétiques, des méthodes de production de polypeptides tels que définis ci-avant, ainsi que des méthodes de criblage ou de caractérisation de composés actifs.

Comme indiqué ci-avant, un premier aspect de l'invention réside dans un composé capable d'interférer, au moins partiellement, au niveau de l'interaction entre la protéine PAP1 (ou ses homologues) et la parkine.

Au sens de la présente invention, la dénomination protéine PAP1 désigne la protéine en soit ainsi que toutes ses formes homologues. Par forme homologue on entend désigner toutes protéines équivalentes à la protéine considérée, d'origine cellulaire diverse et notamment issue de cellules d'origine humaine, ou d'autres organismes, et possédant une activité de même type. De tels homologues comprennent également les variants naturels de la protéine PAP1 de séquence SEQ ID NO2, notamment les variants polymorphiques ou d'épissage. De tels homologues peuvent être obtenus par des expériences d'hybridation entre les acides nucléiques codant (notamment l'acide nucléique de séquence SEQ ID NO:1). Au sens de l'invention, il suffit qu'une séquence de ce type présente un pourcentage d'identité significatif pour conduire à un comportement physiologique assimilable à celui de la protéine PAP1 tel que revendiqué. On entend par pourcentage d'identité significatif un pourcentage d'au moins 60%, préférentiellement de 80%, plus préférentiellement de 90% et encore plus préférentiellement de 95%. A cet égard, des variants et/ou homologues de la séquence SEQ ID NO: 2 sont décrits dans les séquences SEQ ID NO: 13, 15, 43 et 45, identifiés à partir de tissus d'origine humaine. La dénomination PAP1 englobe donc également ces polypeptides.

Le " pourcentage d'identité " entre deux séquences de nucléotides ou d'acides aminés, au sens de la présente invention, peut être déterminé en comparant deux séquences alignées de manière optimale, à travers une fenêtre de comparaison.

La partie de la séquence nucléotidique ou polypeptidique dans la fenêtre de comparaison peut ainsi comprendre des additions ou des délétions (par exemple des " gaps ") par rapport à la séquence de référence (qui ne comprend pas ces additions ou ces délétions) de manière à obtenir un alignement optimal des deux séquences.

Le pourcentage est calculé en déterminant le nombre de positions auquel une base nucléique ou un résidu d'acide aminé identique est observé pour les deux séquences (nucléique ou peptidique) comparées, puis en divisant le nombre de positions auquel il y a identité entre les deux bases ou résidus d'acides aminés par le nombre total de positions dans la fenêtre de comparaison, puis en multipliant le résultat par 100 afin d'obtenir le pourcentage d'identité de séquence.

L'alignement optimal des séquences pour la comparaison peut être réalisé de manière informatique à l'aide d'algorithmes connus contenus dans le package de la Société WISCONSIN GENETICS SOFTWARE PACKAGE, GENETICS COMPUTER GROUP (GCG), 575 Science Doctor, Madison, WISCONSIN.

À titre d'illustration, le pourcentage d'identité de séquence pourra être effectué à l'aide du logiciel BLAST (versions BLAST 1.4.9 de mars 1996, BLAST 2.0.4 de février 1998 et BLAST 2.0.6 de septembre 1998), en utilisant exclusivement les paramètres par défaut (Altschul et al, *J. Mol. Biol.* (1990) 215 : 403-410 ; Altschul et al. *Nucleic Acids Res.* (1997) 25 : 3389-3402). Blast recherche des séquences similaires/homologues à une séquence " requête " de référence, à l'aide de l'algorithme d'Altschul et al. (Supra). La séquence requête et les bases de données utilisées peuvent être peptidiques ou nucléiques, toute combinaison étant possible.

L'interférence d'un composé selon l'invention peut se manifester sous différents aspects. Ainsi, le composé peut ralentir, inhiber ou stimuler, au moins partiellement, l'interaction entre la protéine PAP1 ou l'une de ses formes homologues et la Parkinc. Il s'agit préférentiellement de composés capables de moduler cette

interaction in vitro, par exemple dans un système de type double-hybride ou dans tout système acellulaire de détection d'une interaction entre deux polypeptides. Les composés selon l'invention sont préférentiellement des composés capables de moduler au moins partiellement cette interaction, de préférence d'augmenter ou d'inhiber cette interaction de 20% au moins, plus préférentiellement de 50% au moins, par rapport à un contrôle en absence du composé.

Dans un mode particulier de mise en œuvre, il s'agit de composés capables d'interférer au niveau de l'interaction entre la région de la parkine représentée sur la séquence SEQ ID NO :4 et la région de la protéine PAP1 représentée sur la séquence SEQ ID NO :2, 13, 15, 43 ou 45.

Selon un mode particulier de l'invention, les composés sont capables de se lier au niveau du domaine d'interaction entre la protéine PAP1, ou l'une de ses formes homologues, et la parkine.

Les composés selon la présente invention peuvent être de nature et d'origine variées. En particulier, il peut s'agir de composés de type peptidique, nucléique (i.e., comprenant un enchaînement de bases, notamment une molécule d'ADN ou d'ARN), lipidique, saccharidique, d'un anticorps, etc. et, plus généralement, de toute molécule organique ou inorganique.

Selon une première variante, les composés de l'invention sont de nature peptidique. Le terme peptidique désigne toute molécule comprenant un enchaînement d'acides aminés, tel que par exemple un peptide, un polypeptide, une protéine, un anticorps (ou fragment ou dérivé d'anticorps), le cas échéant modifié ou associé à d'autres composés ou groupements chimiques. A cet égard, le terme « peptide » désigne plus spécifiquement une molécule comprenant un enchaînement de 50 acides aminés au plus, plus préférentiellement de 40 acides aminés au plus. Un polypeptide (ou une protéine) comprend préférentiellement de 50 à 500 acides aminés, ou plus.

Selon un premier mode de mise en œuvre préféré, les composés de l'invention sont des composés peptidiques comprenant tout ou partie de la séquence peptidique SEQ ID N°2 ou un de ses dérivés, en particulier tout ou partie de la séquence peptidique SEQ ID NO : 13, 15, 43 ou 45 ou de dérivés de celles-ci, plus

particulièrement de la protéine PAP1 comprenant la séquence SEQ ID N°2, 13, 15, 43 ou 45.

Au sens de la présente invention, le terme dérivé désigne toute séquence  
différant de la séquence considérée en raison d'une dégénérescence du code génétique,  
5 obtenue par une ou plusieurs modifications de nature génétique et/ou chimique, ainsi  
que tout peptide codé par une séquence hybridant avec la séquence nucléique SEQ ID  
NO : 1 ou un fragment de celle-ci, par exemple avec la séquence nucléique SEQ ID  
NO : 12, 14, 42 ou 44 ou un fragment de celles-ci, et présentant la capacité  
d'interférer au niveau de l'interaction entre la protéine PAP1, ou l'un de ses  
10 homologues, et la Parkine. Par modification de nature génétique et/ou chimique, on  
peut entendre toute mutation, substitution, délétion, addition et/ou modification d'un  
ou plusieurs résidus. Le terme dérivé comprend également les séquences homologues  
à la séquence considérée, issues d'autres sources cellulaires et notamment de cellules  
d'origine humaine, ou d'autres organismes, et possédant une activité de même type.  
15 De telles séquences homologues peuvent être obtenues par des expériences  
d'hybridation. Les hybridations peuvent être réalisées à partir de banques d'acides  
nucléiques, en utilisant comme sonde la séquence native ou un fragment de celle-ci,  
dans des conditions variables d'hybridation (Maniatis *et al.*, 1989). Par ailleurs, le  
terme « fragment » ou « partie » désigne toute portion de la molécule considérée,  
20 comprenant au moins 5 résidus consécutifs, de préférence au moins 9 résidus  
consécutifs, encore plus préférentiellement au moins 15 résidus consécutifs. Des  
fragments typiques peuvent comprendre au moins 25 résidus consécutifs.

De tels dérivés ou fragments peuvent être générés dans des buts différents, tels  
25 que notamment celui d'augmenter leur efficacité thérapeutique ou de réduire leurs  
effets secondaires, ou celui de leur conférer de nouvelles propriétés  
pharmacocinétiques et/ou biologiques.

En tant que peptide dérivé de la protéine PAP1 et des formes homologues, on  
peut citer notamment tout peptide capable d'interagir avec la Parkine, mais portant  
30 une région effectrice rendue non fonctionnelle. De tels peptides peuvent être obtenus



par délétion, mutation ou disruption de cette région effectrice sur la protéine PAP1 et des formes homologues. De telles modifications peuvent être effectuées par exemple par mutagenèse *in vitro*, par introduction d'éléments additionnels ou de séquences synthétiques, ou par des délétions ou des substitutions des éléments originaux.

- 5 Lorsqu'un dérivé tel que défini ci-dessus est réalisé, son activité d'inhibiteur partiel de la fixation de la protéine PAP1 et des formes homologues sur son site de fixation sur la Parkine peut être mise en évidence. Toute technique connue de l'homme du art peut bien évidemment être utilisée à cet effet.

- 10 Il peut également s'agir de fragments des séquences indiquées ci-dessus. De tels fragments peuvent être générés de différentes façons. En particulier, ils peuvent être synthétisés par voie chimique, sur la base des séquences données dans la présente demande, en utilisant les synthétiseurs peptidiques connus de l'homme du métier. Ils peuvent également être synthétisés par voie génétique, par expression dans un hôte
- 15 cellulaire d'une séquence nucléotidique codant pour le peptide recherché. Dans ce cas, la séquence nucléotidique peut être préparée chimiquement en utilisant un synthétiseur d'oligonucléotides, sur la base de la séquence peptidique donnée dans la présente demande et du code génétique. La séquence nucléotidique peut également être préparée à partir des séquences données dans la présente demande, par coupures
- 20 enzymatiques, ligature, clonage, etc, selon les techniques connues de l'homme du métier, ou par criblage de banques d'ADN avec des sondes élaborées à partir de ces séquences.

- 25 Par ailleurs, les peptides de l'invention à savoir capables de moduler au moins partiellement l'interaction entre la protéine PAP1 et des formes homologues et la Parkine peuvent également être des peptides ayant une séquence correspondant au site d'interaction de la protéine PAP1 et des formes homologues sur la Parkine.

- 30 D'autres peptides selon l'invention sont les peptides capables d'entrer en compétition avec les peptides définis ci-dessus pour l'interaction avec leur cible

cellulaire. De tels peptides peuvent être synthétisés notamment sur la base de la séquence du peptide considéré, et leur capacité à entrer en compétition avec les peptides définis ci-dessus peut être déterminée.

5 Un objet spécifique de la présente invention concerne la protéine PAP1. Il s'agit plus particulièrement de la protéine PAP1 comprenant la séquence SEQ ID NO :2 ou un fragment ou dérivé de celle-ci, par exemple la protéine PAP1 de séquence SEQ ID NO : 13, 15, 43, 45 ou de fragments de celles-ci.

10 Un autre objet de l'invention réside dans des anticorps ou fragments ou dérivés d'anticorps polyclonaux ou monoclonaux dirigés contre un polypeptide tel que défini ci-avant. De tels anticorps peuvent être générés par des méthodes connues de l'homme du métier. En particulier, ces anticorps peuvent être préparés par immunisation d'un animal contre un composé peptidique de l'invention (notamment un polypeptide ou  
15 un peptide comprenant tout ou partie de la séquence SEQ ID NO :2), prélèvement du sang, et isolement des anticorps. Ces anticorps peuvent également être générés par préparation d'hybridomes selon les techniques connues de l'homme de l'art.

20 Plus préférentiellement, les anticorps ou fragments d'anticorps de l'invention présentent la capacité de moduler au moins partiellement l'interaction des peptides revendiqués avec la Parkine.

Par ailleurs, ces anticorps peuvent également être utilisés pour détecter et/ou doser l'expression de la PAP1 dans des échantillons biologiques, et de ce fait, pour renseigner sur son état d'activation.

25 Les fragments ou dérivés d'anticorps sont par exemple des fragments Fab, l'ab'2, des anticorps simple-chaîne (ScFv), etc. Il s'agit notamment de tout fragment ou dérivé conservant la spécificité antigénique des anticorps dont ils sont issus.

30 Les anticorps selon l'invention sont plus préférentiellement capables de lier les protéines PAP1 comprenant la séquence SEQ ID NO :2, 13, 15, 43 ou 45, notamment

la région de cette protéine impliquée dans l'interaction avec la parkine. Ces anticorps (ou fragments ou dérivés) sont plus préférentiellement capables de lier un épitope présent dans la séquence comprise entre les résidus 1 et 344 de la séquence SEQ ID NO :2.

5

L'invention concerne également les composés non peptidiques ou non exclusivement peptidiques utilisables comme agent pharmaceutique. Il est en effet possible, à partir des motifs protéiques actifs décrits dans la présente demande, de réaliser des molécules modulatrices de l'activité de PAPI non exclusivement  
10 peptidiques et compatibles avec une utilisation pharmaceutique, en particulier un dupliquant les motifs actifs des peptides avec une structure non peptidique ou de nature non exclusivement peptidique.

La présente invention a également pour objet tout acide nucléique  
15 codant pour un composé peptidique selon l'invention. Il peut s'agir en particulier d'un acide nucléique comprenant tout ou partie de la séquence SEQ ID N°1, 12, 14, 42 ou 44, ou un de ses dérivés. Par séquence dérivée, on entend au sens de la présente invention toute séquence hybridant avec la séquence présentée en SEQ ID N°1 ou avec un fragment de celle-ci et codant pour un composé peptidique selon l'invention,  
20 ainsi que les séquences résultant de ces dernières par dégradation du code génétique. Des acides nucléiques selon l'invention comportent par exemple tout ou partie de la séquence nucléique SEQ ID NO : 12, 14, 42 ou 44.

La présente invention est en outre relative à des séquences présentant un pourcentage d'identité significatif avec la séquence présentée en SEQ ID N°1 ou avec  
25 un fragment de celle-ci et codant pour un composé peptidique présentant un comportement physiologique assimilable à celui de la protéine PAPI. On entend par pourcentage d'identité significatif un pourcentage d'au moins 60%, préférentiellement de 80%, plus préférentiellement de 90% et encore plus préférentiellement de 95%.

Les différentes séquences nucléotidiques de l'invention peuvent être d'origine  
30 artificielle ou non. Il peut s'agir de séquences génomiques, d'ADNc, d'ARN, de

séquences hybrides ou de séquences synthétiques ou semi-synthétiques. Ces séquences peuvent être obtenues soit par criblage de banques d'ADN (banque d'ADNc, banque d'ADN génomique), soit par synthèse chimique, soit par des méthodes mixtes incluant la modification chimique ou enzymatique de séquences obtenues par criblage de banques, soit par recherche d'homologie dans des bases de données nucléiques ou protéiques. L'hybridation mentionnée ci-dessus est préférentiellement réalisée dans les conditions décrites par Sambrook et al (1989, pages 9.52-9.55).

Elle est avantageusement réalisée dans des conditions d'hybridation de forte stringence. Par "conditions d'hybridation de forte stringence" au sens de la présente invention, on entendra les conditions suivantes :

#### 1- Compétition des membranes et PRE HYBRIDATION :

- Mélanger : 40 µl ADN sperme de saumon (10mg/ml)  
+ 40 µl ADN placentaire humain (10mg/ml)
- Dénaturer 5 mn à 96°C, puis plonger dans la glace le mélange.
- Oter le tampon SSC 2X et verser 4 ml de mix formamide dans le tube d'hybridation contenant les membranes.
- Ajouter le mélange des deux ADNs dénaturés.
- Incubation à 42°C pendant 5 à 6 heures, avec rotation.

#### 2- Compétition de la sonde marquée :

- Ajouter à la sonde marquée et purifiée 10 à 50 µl ADN C'ot 1, selon la quantité d'hybridations non spécifiques.
- Dénaturer 7 à 10 mn à 95°C.
- Incuber à 65°C pendant 2 à 5 heures.

#### 3- Hybridation:

- Oter le mix de pré hybridation.
- Mélanger 40 µl ADN sperme de saumon + 40 µl ADN placentaire humain ; dénaturer 5 mn à 96°C, puis plonger dans la glace.

- Ajouter dans le tube d'hybridation 4 ml de mix formamide, le mélange des deux ADN et la sonde marquée/ADN Cot I dénaturée.

- Incuber 15 à 20 heures à 42°C, avec rotation.

#### 4- Lavages :

5 - Un lavage à température ambiante dans du SSC 2X, pour rincer.

- 2 fois 5 minutes à température ambiante SSC 2X et SDS 0,1%.

- 2 fois 15 minutes SSC 0,1X et SDS 0,1% à 65°C.

Envelopper les membranes dans du Saran et exposer.

10 Les conditions d'hybridation décrites plus haut sont adaptées à l'hybridation dans des conditions de forte stringence, d'une molécule d'acide nucléique d'une longueur variable de 20 nucléotides à plusieurs centaines de nucléotides.

Il va sans dire que les conditions d'hybridation ci-dessus décrites peuvent être adaptées en fonction de la longueur de l'acide nucléique dont l'hybridation est recherchée ou du type de marquage choisi, selon des techniques connues de l'homme

15 du métier.

Les conditions convenables d'hybridation peuvent par exemple être adaptées selon l'enseignement contenu dans l'ouvrage de HAMES et HIGGINS (1985) (Nucleic acid Hybridization iapactical Approach, Hames and Higgins Ed., IRL Press, Oxford) ou encore dans l'ouvrage de F. AUSUBEL et al (1999) (Currents Protocols in Molecular Biology, Green Publishing Associates and Wiley Interscience, N.Y).

20 Un acide nucléique particulier au sens de l'invention code pour un polypeptide comprenant la séquence SEQ ID NO :2 ou un fragment ou dérivé de celle-ci, notamment pour la protéine PAP1 humaine. Il s'agit avantageusement d'un acide

25 nucléique comprenant la séquence nucléique SEQ ID NO :1, 12, 14, 42 ou 44.

De tels acides nucléiques peuvent être utilisés pour la production des composés peptidiques de l'invention. La présente demande concerne ainsi un procédé de préparation de tels composés peptidiques selon lequel on cultive une cellule

30 contenant un acide nucléique selon l'invention, dans des conditions d'expression dudit

acide nucléique et on récupère le composé peptidique produit. Dans ce cas, la particodant pour ledit composé peptidique est généralement placée sous le contrôle de signaux permettant son expression dans un hôte cellulaire. Le choix de ces signaux (promoteurs, terminateurs, séquence "leader" de sécrétion, etc) peut varier en fonction de l'hôte cellulaire utilisé. Par ailleurs, les acides nucléiques de l'invention peuvent faire partie d'un vecteur qui peut être à répllication autonome ou intégratif. Plus particulièrement, des vecteurs à répllication autonome peuvent être préparés en utilisant des séquences à répllication autonome chez l'hôte choisi. S'agissant des vecteurs intégratifs, ceux-ci peuvent être préparés par exemple en utilisant des séquences homologues à certaines régions du génome de l'hôte, permettant, par recombinaison homologue, l'intégration du vecteur. Il peut s'agir d'un vecteur de type plasmidique, épisomique, chromosomique, viral, etc.

Les hôtes cellulaires utilisables pour la production des composés peptidiques de l'invention par voie recombinante sont aussi bien des hôtes eucaryotes que procaryotes. Parmi les hôtes eucaryotes qui conviennent, on peut citer les cellules animales, les levures, ou les champignons. En particulier, s'agissant de levures, on peut citer les levures du genre *Saccharomyces*, *Kluyveromyces*, *Pichia*, *Schwannomyces*, ou *Hansenula*. S'agissant de cellules animales, on peut citer les cellules COS, CHO, C127, PC12 etc. Parmi les champignons, on peut citer plus particulièrement *Aspergillus* ssp. ou *Trichoderma* ssp. Comme hôtes procaryotes, on préfère utiliser les bactéries suivantes *E. coli*, *Bacillus*, ou *Streptomyces*.

La présente invention a en outre pour objet des mammifères non-humains comprenant dans leurs cellules un acide nucléique ou un vecteur selon l'invention.

De tels mammifères (rongeurs, canins, lapins, etc.) sont notamment utilisables pour l'étude des propriétés de la PAP1 et l'identification de composés à visée thérapeutique. La modification du génome d'un tel animal transgénique peut résulter d'une altération ou une modification de un ou plusieurs gènes par "knock-in" ou par "knock-out". Cette modification peuvent être effectuées à l'aide d'agents altérants ou mutagènes classiques ou bien par mutagenèse dirigée. La modification du génome peut également résulter d'une insertion de gène(s) ou de remplacement de gène(s)

dans sa (leur) forme sauvage ou mutée. Les modifications du génome sont avantageusement effectuées sur des cellules souches reproductrices et avantageusement sur les pronucléi. La transgénèse peut être effectuée par microinjection d'une cassette d'expression comprenant les gènes modifiés dans les deux pronucléi fécondés. Ainsi un animal selon la présente invention peut être obtenu par injection d'une cassette d'expression comprenant un acide nucléique. De manière préférentielle, cet acide nucléique est un ADN qui peut être un ADN génomique (ADNg) ou un ADN complémentaire (ADNc).

La construction d'animaux transgéniques selon l'invention peut être réalisée selon des techniques classiques bien connues de l'homme du métier. L'homme du métier pourra en particulier se référer à la production d'animaux transgéniques, et particulièrement à la production de souris transgéniques, telles que décrites dans les brevets US 4,873,191, US 5,464,764 et US 5,789,215. Le contenu de ces documents étant ici incorporé par référence.

En bref, une construction polynucléotidique comprenant un acide nucléique selon l'invention est insérée dans une lignée de cellules souches de type ES. L'insertion de la construction polynucléotidique est réalisée de préférence par électroporation, telle que décrite par Thomas et al. (1987, Cell, Vol.51:503-512).

Les cellules ayant subi l'étape d'électroporation sont ensuite criblées pour la présence de la construction polynucléotidique (par exemple par sélection à l'aide de marqueurs, ou encore par PCR ou par analyse sur gel d'électrophorèse d'ADN de type Southern) afin de sélectionner les cellules positives ayant intégré la construction polynucléotidique exogène dans leur génome, le cas échéant à la suite d'un événement de recombinaison homologue. Une telle technique est par exemple décrite par MANSOUR et al. (*Nature* (1988) 336: 348-352).

Ensuite, les cellules sélectionnées positivement sont isolées, clonées et injectées dans des blastocystes de souris de 3,5 jours, comme cela est décrit par BRADLEY (1987, *Production and Analysis of Chimaeric mice*. In: E. J. ROBERTSON (Ed., *teratocarcinomas and embryonic stem cells: A practical approach*. IRL press, Oxford, page 113). Des blastocystes sont ensuite introduits dans

un animal hôte femelle et le développement de l'embryon est poursuivi jusqu'au terme.

Selon une alternative, des cellules de type ES sélectionnées positivement sont mises en contact avec des embryons de 2.5 jours à un stade 8-16 cellules (*morulae*) comme décrit par WOOD et al. (1993, Proc. Natl. Acad. Sci. USA, vol.90: 4582-4585) ou par NAGY et al. (1993, Proc. Natl. Acad. Sci. USA, vol. 90: 8424-8428), les cellules ES étant internalisées afin de coloniser extensivement le blastocyste, y compris les cellules donnant naissance à la lignée germinale.

Les descendants sont ensuite testés afin de déterminer ceux qui ont intégré la construction polynucléotidique (le transgène).

Les acides nucléiques selon l'invention peuvent également servir à la réalisation d'oligonucléotides antisens ou d'antisens génétiques utilisables comme agents pharmaceutiques. Les séquences antisens sont des oligonucléotides de petite taille, complémentaire du brin codant d'un gène donné, et de ce fait capables d'hybrider spécifiquement avec l'ARNm transcrit, inhibant sa traduction en protéine. L'invention a ainsi pour objet les séquences antisens capables d'inhiber au moins partiellement l'interaction des protéines PAP1 sur la Parkine. De telles séquences peuvent être constituées par tout ou partie des séquences nucléiques définies ci-avant. Il s'agit généralement de séquences ou de fragments de séquences complémentaires de séquences codant pour des peptides interagissant avec la Parkine. De tels oligonucléotides peuvent être obtenus par fragmentation, etc., ou par synthèse chimique.

Les séquences revendiquées peuvent être utilisées dans le cadre de thérapies géniques, pour le transfert et l'expression *in vivo* de séquences antisens ou de peptides capables de moduler l'interaction de la protéine PAP1 avec la Parkine. A cet égard, les séquences peuvent être incorporées dans des vecteurs viraux ou non viraux, permettant leur administration *in vivo* (Kahn *et al.*, 1991). A titre de vecteurs viraux conformes à l'invention on peut tout particulièrement citer les vecteurs de type



adénovirus, rétrovirus, virus associé à l'adénovirus (AAV) ou virus de l'herpès. La présente demande a également pour objet des virus recombinants défectifs comprenant un acide nucléique codant pour un polypeptide selon l'invention, notamment un polypeptide ou peptide comprenant tout ou partie de la séquence SEQ ID NO : 2 ou d'un dérivé de celle-ci, par exemple tout ou partie de la séquence SEQ ID NO : 12, 14, 42 ou 44 ou de dérivées de celles-ci.

L'invention permet également la réalisation de sondes nucléotidiques, synthétiques ou non, capables de s'hybrider avec les séquences nucléotidiques définies ci-avant ou leur brin complémentaire. De telles sondes peuvent être utilisées *in vitro* comme outil de diagnostic, pour la détection de l'expression ou surexpression de PAPI, ou encore pour la mise en évidence d'anomalies génétiques (mauvais épissage, polymorphisme, mutations ponctuelles, etc). Ces sondes peuvent également être utilisées pour la mise en évidence et l'isolement de séquences d'acides nucléiques homologues codant pour des peptides tels que définis précédemment, à partir d'autres sources cellulaires et préférentiellement de cellules d'origines humaines. Les sondes de l'invention comportent généralement au moins 10 bases, et elles peuvent par exemple comporter jusqu'à l'intégralité d'une des séquences précitées ou de leur brin complémentaire. Préférentiellement, ces sondes sont marquées préalablement à leur utilisation. Pour cela, différentes techniques connues de l'homme du métier peuvent être employées (marquage radioactif, fluorescent, enzymatique, chimique, etc).

L'invention concerne également des amorces ou couples d'amorces permettant d'amplifier tout ou une partie d'un acide nucléique codant une PAPI, par exemple une amorce de séquence choisie parmi SEQ ID NO : 16-41.

L'invention a encore pour objet toute composition pharmaceutique comprenant comme principe actif au moins un composé tel que défini ci-dessus, notamment un composé peptidique.

Elle a notamment pour objet toute composition pharmaceutique comprenant comme principe actif au moins un anticorps et/ou un fragment d'anticorps tel que

défini ci-dessus, ainsi que toute composition pharmaceutique comprenant comme principe actif au moins un acide nucléique ou un vecteur tel que défini ci-dessus.

Elle a également pour objet toute composition pharmaceutique comprenant comme principe actif une molécule chimique capable d'augmenter ou de diminuer l'interaction entre la protéine PAPI et la Parkine.

Par ailleurs, elle a aussi pour objet les compositions pharmaceutiques dans lesquelles les peptides, anticorps, molécules chimiques et séquences nucléotidiques définis ci-dessus sont associés entre-eux ou avec d'autres principes actifs.

- 10 Les compositions pharmaceutiques selon l'invention peuvent être utilisées pour moduler l'activité de la protéine Parkine et de ce fait maintenir la survie des neurones dopaminergiques. Plus particulièrement, ces compositions pharmaceutiques sont destinées à moduler l'interaction entre la protéine PAPI et la Parkine. Il s'agit plus préférentiellement de compositions pharmaceutiques destinées au traitement de
- 15 maladies du système nerveux central comme par exemple la maladie de Parkinson.

- L'invention a encore pour objet l'utilisation des molécules décrites ci-avant pour moduler l'activité de la Parkine ou le typage de maladies du système nerveux central. En particulier, l'invention concerne l'utilisation de ces molécules pour
- 20 moduler au moins partiellement l'activité de la Parkine.

- L'invention concerne également un procédé pour le criblage ou la caractérisation de molécules actives sur la fonction de la parkine, comprenant la sélection de molécules capables de lier la séquence SEQ ID NO :2 ou la séquence
- 25 SEQ ID NO :4, ou un fragment (ou dérivé) de celles-ci. Le procédé comprend avantageusement la mise en contact, in vitro, de la ou des molécules test avec un polypeptide comprenant la séquence SEQ ID NO :2 ou la séquence SEQ ID NO :4, ou un fragment (ou dérivé) de celles-ci, et la sélection de molécules capables de lier la séquence SEQ ID NO :2 (notamment la région comprise entre les résidus 1 et 344) ou
- 30 la séquence SEQ ID NO :4. Les molécules testées peuvent être de nature variée

(peptide, nucléique, lipide, sucre, etc., ou des mélanges de telles molécules, par exemple des bibliothèques combinatoires, etc.). Comme indiqué ci-avant, les molécules ainsi identifiées peuvent être utilisées pour moduler l'activité de la protéine Parkine, et représentent des agents thérapeutiques potentiels pour le traitement de pathologies neurodégénératives.

D'autres avantages de la présente invention apparaîtront à la lecture des exemples et figure qui suivent, qui doivent être considérés comme illustratifs et non limitatifs.

#### LEGENDES DE LA FIGURE :

Figure 1: Représentation du vecteur pLex9-Parkine (135-290)

Figure 2 : Résultats de la première expérience 5'-RACE. 8 clones ont été obtenus. La séquence électronique initiale est indiquée au bas de la figure.

Figure 3 : Résultats de la seconde expérience 5'-RACE. Seulement deux des 8 clones obtenus dans la première expérience ont été validés (clones A12 et D5). La séquence électronique initiale est indiquée au bas de la figure. La séquence complète des ADN et protéines est donnée dans les Séquences 12-15.

Figure 4 : Détail de l'organisation des clones C5 et D4 de la seconde expérience 5'-RACE. La séquence consensus résultante est indiquée en haut de la figure.

Figure 5 : Structure des transcrits isolés à partir de cerveau humain.

Figure 6 : Séquence nucléique et protéique de LY111 (full length) de cerveau humain. Double souligné : cystéines conservées du domaine en doigt de zinc. Gras : Domaine C<sub>2</sub>1. Italique : domaine C<sub>2</sub>2.

Figure 7 : Séquence nucléique et protéique de LY111 (version courte) de cerveau humain. Double souligné : cystéines conservées du domaine en doigt de zinc. Gras : Domaine C<sub>2</sub>1, Italique : domaine C<sub>2</sub>2.

Figure 8 : Localisation de la protéine LY111 courte (8b) ou longue (8a) après expression dans les cellules Cos-7.

Figure 9 : Séquence nucléique et protéique de LY111 (version longue) de poumon humain.

Figure 10 : Séquence nucléique et protéique de LY111 (version courte) de cerveau humain.

5

## MATERIELS ET TECHNIQUES MIS EN ŒUVRE

### 1) Souches de levure:

- La souche L40 du genre *S.cerevisiae* (Mata, *his3D200*, *trp1-901*, *leu2-3,112*,  
 10 *ade2*, *LYS2:: (lexAop)<sub>r</sub>-HIS3*, *URA3:: (lexAop)<sub>r</sub>-LacZ*, *GAL4*, *GAL80*) a été utilisée pour vérifier les interactions protéine-protéine quand l'un des partenaires protéiques est fusionné à la protéine LexA. Cette dernière est capable de reconnaître l'élément de réponse LexA contrôlant l'expression des gènes rapporteurs *LacZ* et *HIS3*.

Elle a été cultivée sur les milieux de culture suivants:

- 15 Milieu YPD complet : -Extrait de levures (10 g/l) (Difco)  
 -Bactopeptone (20 g/l) (Difco)  
 -Glucose (20 g/l) (Merck)

Ce milieu a été rendu solide par addition de 20 g/l d'agar (Difco).

- 20 Milieu YNB minimum : -Yeast Nitrogen Base (sans acides aminés) (6,7 g/l) (Difco)  
 - Glucose (20 g/l) (Merck)

Ce milieu peut être rendu solide par addition de 20 g/l d'agar (Difco). Il peut également être supplémenté en acides aminés et/ou en 3-amino-1,2,4-triazole par  
 25 addition de milieux CSM [CSM -Leu, -Trp, -His (620 mg/l), CSM -Trp (740 mg/l) ou CSM -Leu, -Trp (640 mg/l)(Bio101)] et/ou de 3-amino-1,2,4-triazole 2,5 mM.

### 2) Souches de bactéries:

- La souche TG1 d'*Escherichia coli*, de génotype supE, *hsdΔ5*, *thi*,  $\Delta(lac$ -  
 30  $proAB)$ , F'[*tra* D36 *pro* A<sup>+</sup>B<sup>+</sup> *lacI*<sup>d</sup> *lacZΔM15*], a été employée pour la construction

de plasmides, comme moyen d'amplification et d'isolement de plasmides recombinants utilisés. Elle a été cultivée sur le milieu suivant :

Milieu LB:

- NaCl (5 g/l) (Prolabo)
- Bactotryptone (10 g/l) (Difco)
- Extrait de levure (5 g/l) (Difco)

Ce milieu est rendu solide par addition de 15 g/l d'agar (Difco).

L'ampicilline a été utilisée à 100 µg/ml, cet antibiotique sert à sélectionner les bactéries ayant reçu les plasmides portant comme marqueur le gène de résistance à cet antibiotique.

- 10 La souche HB101 d'*Escherichia coli* de génotype supE44, ara14, galK2, lacY1,  $\Delta$ (gpt-prfA)62, rpsL20(Str<sup>r</sup>), xyl-5, mtl-1, recA13,  $\Delta$ (mcrC-mrr), HsdS<sup>-</sup>(r<sup>-</sup>m<sup>-</sup>) a été employée comme moyen d'amplification et d'isolement de plasmides provenant de la banque d'ADNc de lymphocyte humain.

Elle a été cultivée sur

- 15 Milieu M9:
- Na2HPO4 (7 g/l) (Prolabo)
  - K112PO4 (3 g/l) (Prolabo)
  - NH4Cl (1 g/l) (Prolabo)
  - NaCl (0,5 g/l) (Prolabo)
  - Glucose (20 g/l) (Sigma)
  - 20 -MgSO4 (1 mM) (Prolabo)
  - Thiamine (0,001 %) (Sigma)

Ce milieu est rendu solide par addition de 15 g/l d'agar (Difco).

De la leucine (50 mg/l) (Sigma) et de la proline (50 mg/l) (Sigma) doivent être ajoutées au milieu M9 pour permettre la croissance de la souche HB101.

- 25 Lors de la sélection de plasmides provenant de la banque deux hybrides d'ADNc de lymphocyte, la leucine n'a pas été ajoutée au milieu car les plasmides portent un marqueur de sélection Leu2.

### 3) Plasmides :

Le vecteur pLex9 (pBTM116) (Bartel *et al.*, 1993) de 5kb homologue au pGBT10 qui contient un site multiple de clonage situé en aval de la séquence codant pour le répresseur bactérien LexA et en amont d'un terminateur pour former une protéine de fusion.

- 5 pLex-HaRasVal12, plasmide pLex9, tel que décrit dans la demande WO98/21327, qui contient la séquence codant pour la protéine HaRas mutée en position Val12 connue pour interagir avec la protéine Raf de mammifère (Vojtek *et al.*, 1993). Ce plasmide a été utilisé pour tester la spécificité d'interaction de la protéine PAP1 dans la souche L40.

- 10 pLex9-cAPP, plasmide pLex9 qui contient la séquence codant pour le domaine cytoplasmique de la protéine APP connue pour interagir avec le domaine PTB2 de FE65. Ce plasmide a été utilisé pour tester la spécificité d'interaction de la protéine PAP1 dans la souche L40.

15 **4) Oligonucléotides de synthèse:**

TTAAGAATTC GGAAGTCCAG CAGGTAG (SEQ ID N°5)

ATTAGGATCC CTACACACAA GGCAGGGAG (SEQ ID N°6)

Oligonucléotides qui ont permis d'obtenir le fragment PCR correspondant à la région centrale de la Parkine bordée par les sites *EcoRI* et *BamHI*.

20

GCGTTTGGA TCACTACAG (SEQ ID N°7)

GGTCTCGGT TGGCATC (SEQ ID N°8)

CCGCTTGCTT GGAGGAAC (SEQ ID N°9)

CGTATTTCTC CGCCTTGG (SEQ ID N°10)

- 25 AATAGCTCGA GTCAGTGCA GACAAGAG (SEQ ID N°11)

Oligonucléotides qui ont servi à séquencer l'insert correspondant au gène PAP1.

Les oligonucléotides sont synthétisés sur l'appareil Applied System ABI 394-08. Ils sont décrochés de la matrice de synthèse par de l'ammoniac et précipités deux

fois par 10 volumes de n-butanol puis repris dans de l'eau. La quantification est effectuée par mesure de la densité optique ( $1\text{DO}_{260}$  correspond à  $30\text{ }\mu\text{g/ml}$ ).

### 5) Préparation des ADN plasmidiques.

- 5 Les préparations en petite quantité et en grande quantité d'ADN plasmidique ont été effectuées selon les protocoles recommandés par le fabricant Quiagen des kits de purification d'ADN :

- Quiaprep Spin Miniprep kit, ref : 27106
- Quiaprep Plasmid Maxiprep kit, ref : 12163.

10

### 6) Amplification enzymatique d'ADN par PCR (Polymerase Chain Reaction) :

- Les réactions de PCR sont effectuées dans un volume final de  $100\mu\text{l}$  en présence de la matrice d'ADN, de dNTP ( $0,2\text{ mM}$ ), de tampon PCR (Tris-HCl pH 8,5  $10\text{ mM}$ ,  $\text{MgCl}_2$   $1\text{ mM}$ , KCl  $5\text{ mM}$ , gélatine  $0,01\%$ ), de  $10\text{-}20\text{ pmoles}$  de chacun des
- 15 oligonucléotides et de  $2,5\text{ UI}$  d'Ampli Taq DNA polymérase (Perkin Elmer). Le mélange est recouvert de 2 gouttes d'huile de paraffine pour limiter l'évaporation de l'échantillon. L'appareil utilisé est le "Crocodile II" d'Appligene.

- Nous avons utilisé une température de dénaturation de la matrice de  $94^\circ\text{C}$ , une température d'hybridation de  $52^\circ\text{C}$  et une température d'élongation par l'enzyme à
- 20  $72^\circ\text{C}$ .

### 7) Les ligatures :

- Toutes les réactions de ligation sont effectuées à  $37^\circ\text{C}$  pendant une heure dans un volume final de  $20\text{ }\mu\text{l}$  en présence de  $100\text{ à }200\text{ ng}$  de vecteur,  $0,1\text{ à }0,5\text{ }\mu\text{g}$
- 25 d'insert,  $40\text{ UI}$  d'enzyme T4 DNA ligase (Biolabs) et un tampon de ligation (Tris-HCl  $50\text{ mM}$  pH 7,8;  $\text{MgCl}_2$   $10\text{ mM}$ ; DTT  $10\text{ mM}$ ; ATP  $1\text{ mM}$ ). Le témoin négatif est constitué par la ligation du vecteur en l'absence d'insert.

### 8) Tranformation des bactéries :

La transformation des bactéries par un plasmide est effectuée selon le protocole suivant : 10µl du volume de ligation est utilisée pour transformer les bactéries TG1 selon la méthode la méthode de Chung (Chung *et al.*, 1989). Après transformation les bactéries sont étalées sur un milieu LB + ampicilline et incubées 16h à 37°C.

### 9) Séparation et extraction des ADN:

La séparation des ADN est réalisée en fonction de leur taille par électrophorèse sur gel d'agarose selon Maniatis (Maniatis *et al.*, 1989) : gel d'agarose à 1% (Gibco BRL) dans un tampon TBE (Tris base 90 mM; Borate 90 mM; EDTA 2 mM)

### 10) Séquençage fluorescent des ADN plasmidiques :

La technique de séquençage utilisée est dérivée de la méthode de Sanger (Sanger *et al.*, 1977) et adaptée pour le séquençage par fluorescence développé par Applied Biosystems. Le protocole utilisé est celui décrit par les concepteurs du système (Perkin Elmer, 1997).

### 11) Transformation de la levure:

Les plasmides sont introduits dans la levure par une technique classique de transformation de la levure mise au point par Gietz (Gietz *et al.*, 1992) et modifiée de la façon suivante :

Dans le cas particulier de la transformation de la levure par la banque d'ADNc de lymphocyte, la levure utilisée contient le plasmide pLex9-Parkine (135-290) codant pour la partie centrale de la Parkine fusionnée à la protéine LexA. Elle est cultivée dans 200 ml de milieu minimum YNB supplémenté en acides aminés CSM - Trp à 30°C sous agitation jusqu'à une densité de  $10^7$  cellules/ml. Pour effectuer la transformation des levures suivant le protocole précédant, la suspension cellulaire a été séparée en 10 tubes de 50µl dans lesquels 5µg de la banque ont été ajoutés. Le choc thermique a été effectué pendant 20 minutes puis les cellules ont été collectées par centrifugation et resuspendues dans 100 ml de milieu YPD pendant 1h à 30°C et



dans 100 ml de milieu YNB supplémenté en CSM -Leu, -Trp pendant 3h30 à 30°C. L'efficacité de la transformation est déterminée en étalant différentes dilutions de cellules transformées sur milieu YNB solide supplémenté en CSM -Trp, -Leu. Après une culture à 30°C durant 3 jours les colonies obtenues ont été comptées et le taux de transformation par  $\mu\text{g}$  d'ADN de la banque de lymphocyte a été déterminé.

### 12) Isolement de plasmides extrait de levure :

5 ml d'une culture de levure incubée 16h à 30°C sont centrifugés et repris dans 200  $\mu\text{l}$  d'un tampon de lyse (Sorbitol 1M,  $\text{KH}_2\text{PO}_4/\text{K}_2\text{HPO}_4$  0,1M pH7,4, zymolyase 12,5 mg/ml) et incubés 1h à 37°C. Le lysat est ensuite traité selon le protocole recommandé par le fabricant Quiagen du kit de purification d'ADN, Quiaprep Spin Miniprep kit, ref 27106.

### 13) Test d'activité de la $\beta$ -galactosidase :

Une feuille de nitrocellulose est préalablement déposée sur la boîte de Pétri contenant les clones de levures individualisés. Cette feuille est ensuite plongée dans de l'azote liquide pendant 30 secondes afin de faire éclater les levures et de libérer ainsi l'activité  $\beta$ -galactosidase. Après décongélation, la feuille de nitrocellulose est déposée, colonies vers le haut, dans une autre boîte de Pétri contenant un papier Whatman préalablement imbibé de 1,5 ml de solution PBS ( $\text{Na}_2\text{HPO}_4$  60 mM,  $\text{NaH}_2\text{PO}_4$  40 mM, KCl 10 mM,  $\text{MgSO}_4$  1 mM, pH7) contenant 15  $\mu\text{l}$  de X-Gal (5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactoside) à 40 mg/ml de N,N-diméthylformamide. La boîte est ensuite placée dans une étuve à 37°C. Le test est dit positif lorsque les colonies virent au bleu sur la membrane au bout de 12 heures.

**EXEMPLE 1: CONSTRUCTION D'UN VECTEUR PERMETTANT L'EXPRESSION D'UNE PROTEINE DE FUSION ENTRE LA PARTIE CENTRALE DE LA PARKINE ET LE REPRESSEUR BACTERIEN LEXA.**

Le criblage d'une banque utilisant le système double hybride nécessite que la région centrale de la Parkine soit fusionnée à une protéine liant l'ADN comme le répresseur bactérien LexA. L'expression de cette protéine de fusion est réalisée grâce au vecteur pLex9 (cf. matériels et méthodes), dans lequel a été introduite, dans le même cadre de lecture que la séquence correspondant à la protéine LexA, la séquence codant pour la région centrale de la Parkine figurant dans la séquence présentée en SEQ ID N°3 ou 4.

Le fragment d'ADN de 468 pb correspondant aux 156 acides aminés de la région centrale de la Parkine qui commence à l'acide aminé 135 a été obtenu par PCR à partir des oligonucléotides (SEQ ID N°5 et N°6) qui ont également permis d'introduire le site *EcoRI* à l'extrémité 5' et un codon stop et un site *BamHI* à l'extrémité 3'. Le fragment de PCR a été introduit entre les sites *EcoRI* et *BamHI* du multisite de clonage du plasmide pLex9 en aval de la séquence codant pour la protéine LexA pour donner le vecteur pLex9-Parkine (135-290) (Fig.1).

La construction a été vérifiée par séquençage de l'ADN. Cette vérification a permis de montrer que ce fragment ne présentait pas de mutations générées au cours de la réaction de PCR et qu'il était fusionné dans la même phase ouverte de lecture que celle du fragment correspondant à LexA.

## EXEMPLE 2: CRIBLAGE D'UNE BANQUE DE FUSION DE LYMPHOCYTE.

Nous avons utilisé la méthode du Double-Hybride (Fields et Song, 1989).

Le criblage d'une banque de fusion permet d'identifier des clones produisant des protéines fusionnées au domaine transactivateur de GAL4, pouvant interagir avec la protéine d'intérêt décrite dans l'exemple 1 (région centrale de la parkine). Cette interaction permet de reconstituer un transactivateur qui va alors être capable d'induire l'expression des gènes rapporteurs *His3* et *LacZ* dans la souche L40.

Pour effectuer ce criblage nous avons choisi une banque de fusion réalisée à partir d'ADN provenant de lymphocytes humains périphériques fournie par Richard

Benarous (Peytavi *et al.*, 1999). Des levures ont été transformées par la banque de lymphocytes et des clones positifs ont été sélectionnés comme décrit ci-après.

Lors du criblage il est nécessaire de préserver la probabilité que chaque plasmide indépendant de la banque de fusion soit présent dans au moins une levure en même temps que le plasmide pLex9-Parkine (135-290). Pour préserver cette probabilité il est important d'avoir une bonne efficacité de transformation de la levure. Pour cela nous avons choisi un protocole de transformation de la levure donnant une efficacité de  $2.6 \cdot 10^5$  cellules transformées par  $\mu\text{g}$  d'ADN. De plus comme la cotransformation de la levure par deux plasmides différents réduit cette efficacité, nous avons préféré utiliser une levure préalablement transformée par le plasmide pLex9-Parkine (135-290). Cette souche L40 pLex9-Parkine (135-290) de phénotype His<sup>-</sup> Lys<sup>-</sup> Leu<sup>-</sup> Adé<sup>-</sup> a été transformée par 50 $\mu\text{g}$  d'ADN plasmidique de la banque de fusion. Cette quantité d'ADN nous a permis d'obtenir après estimation  $1.3 \cdot 10^7$  cellules transformées, ce qui correspond à un nombre légèrement supérieur au nombre de plasmides indépendants qui constituent la banque. D'après ce résultat nous pouvons penser que la quasi totalité des plasmides de la banque a servi à transformer les levures. La sélection des cellules transformées, capables de reconstituer un transactivateur fonctionnel, a été faite sur un milieu YNB supplémenté avec 2.5 mM 3-amino-1,2,4-triazole et 620 mg/l de CSM (Bio101) ne contenant pas d'histidine de leucine et de tryptophane.

A l'issue de cette sélection, de nombreux clones avec un phénotype His<sup>+</sup> ont été obtenus. Sur ces transformants un test d'activité  $\beta$ -galactosidase a été effectué afin de valider par l'expression de l'autre gène rapporteur, *LacZ*, ce nombre de clones obtenus. 115 clones présentaient le double phénotype His<sup>+</sup>,  $\beta$ -Gal<sup>+</sup> pouvant correspondre à une interaction protéine-protéine.

### EXEMPLE 3: ISOLEMENT DES PLASMIDES DE LA BANQUE DANS LES CLONES SELECTIONNES.

Afin d'identifier les protéines pouvant interagir avec la région centrale de la Parkine, les plasmides de la banque de fusion contenus dans les levures sélectionnées lors du criblage double hybride ont été extraits. Pour pouvoir en obtenir en grande quantité, cet isolement nécessite au préalable une transformation d'*E.coli* par un extrait d'ADN des souches de levures positives. Comme le plasmide de la banque contenu dans cet extrait est un plasmide navette levure/*E.coli* il peut facilement se répliquer chez la bactérie. Le plasmide de la banque a été sélectionné par complémentatoin de la bactérie HB101 auxotrophe pour la leucine sur milieu dépourvu de leucine.

Les ADN plasmidiques des colonies bactériennes obtenues après transformation par des extraits d'ADN de levures ont été analysés par digestion avec des enzymes de restriction et séparation des fragments d'ADN sur gel d'agarose. Parmi les 115 clones analysés, un clone contenant un plasmide de la banque présentant un profil différent des autres a été obtenu. Ce plasmide, appelé pGAD-Ly111b, a été plus précisément étudié.

#### EXEMPLE 4: DETERMINATION DE LA SEQUENCE DE L'INSERT CONTENU DANS LE PLASMIDE IDENTIFIE.

Le séquençage de l'insert contenu dans le plasmide identifié a été réalisé dans un premier temps à partir de l'oligonucléotide SEQ ID N°7 complémentaire de la séquence de GAL4TA à proximité du site *EcoRI* d'insertion de la banque de cDNA de lymphocytes. Puis dans un deuxième temps à partir des oligonucléotides SEQ ID N°8 à SEQ ID N°11, correspondant à la séquence de l'insert obtenue au cours de la progression du séquençage. La séquence obtenue est présentée sur la séquence SEQ ID NO1. La protéine ainsi identifiée a été désignée PAP1 (Parkin-Associated Protein 1).

La comparaison de la séquence de cet insert avec les séquences contenues dans les banques de données GENBank et EMBL (European Molecular Biology Lab)

a montré une homologie de 25%, au niveau protéique, avec différents membres de la famille des Synaptotagmines. Les Synaptotagmines font partie d'une famille de protéines membranaires codées par au moins onze gènes différents exprimés dans le cerveau et autres tissus. Elles contiennent un domaine transmembranaire unique et deux domaines régulés par le calcium appelés  $C_2$ . C'est dans ce domaine que l'on trouve l'homologie entre les Synaptotagmines et la protéine PAP1. Aucune autre homologie significative n'a été observée.

#### EXEMPLE 5: ANALYSE DE LA SPECIFICITE D'INTERACTION ENTRE LA REGION CENTRALE DE LA PARKINE ET LA PROTEINE PAP1.

Afin de déterminer la spécificité d'interaction entre le fragment correspondant à la protéine PAP1 et la région centrale de la Parkine, un test d'interaction spécifique deux hybrides avec d'autres protéines non pertinentes a été réalisé. Pour réaliser ce test nous avons transformé la souche L40 par les plasmides de contrôle pLex9-cAPP ou pLex9-HaRasVal12 à la place du plasmide pLex9-Parkin (135-290) codant respectivement pour le domaine cytoplasmique de l'APP ou la protéine HaRasVal12 fusionnés au domaine de liaison à l'ADN de LexA et par le plasmide isolé lors du criblage de la banque deux hybrides. Un test d'activité  $\beta$ -Gal sur les cellules transformées par les différents plasmides a été effectué afin de déterminer une interaction protéine-protéine. D'après le résultat du test, seules les levures transformées par le plasmide isolé lors du criblage de la banque deux hybrides et par le plasmide pLex9-Parkin (135-290) présentaient une activité  $\beta$ -Gal+ montrant ainsi une interaction entre la région centrale de la Parkine et la protéine PAP1. Cette interaction s'avère donc être spécifique puisque ce fragment de PAP1 ne semble pas interagir avec les protéines cAPP ou HaRasVal12.

Ces résultats montrent donc l'existence d'une protéine nouvelle, désignée PAP1, capable d'interagir de manière spécifique avec la Parkine. Cette protéine, apparentée aux synaptotagmines, ne présente aucune homologie significative avec des

protéines connues, et peut être utilisée dans des applications thérapeutiques ou diagnostiques, pour la production d'anticorps, de sondes ou peptides ou encore pour le criblage de molécules actives.

## 5 EXEMPLE 6 : CLONAGE DU GENE PAPI A PARTIR D'UNE BANQUE D'ADN DE POU MON HUMAIN

Dans le but d'identifier la séquence complète du gène PAPI humain et de caractériser l'existence de formes variantes, deux approches d'elongation électronique ont été réalisées à partir de la séquence SEQ ID NO :1. Deux séquences électroniques ont ainsi été obtenues, de 1644 pb et 1646 pb respectivement, comportant une elongation de 330 pb par rapport à la séquence SEQ ID NO :1. Néanmoins, l'analyse de ces séquences a montré des différences dans la région consensus, apparentes après traduction. Ainsi, une ORF de 420aa est obtenue dans un cas, et une ORF de 230aa avec l'autre séquence. La séquence protéique obtenue a été comparée avec les séquences connues, et a révélé une homologie de 24% sur les 293 acides aminés recouvrant avec la synaptogamine humaine 1 (p65)(p21579). La fonction de la synaptogamine 1 meut être un rôle régulateur dans les interactions membranaires au cours du trafic de vésicules synaptiques dans la zone de la synapse. Elle lie les phospholipides acides avec une certaine spécificité. En outre, une interaction dépendante du calcium entre la synaptogamine et les récepteurs de protéine kinase C activée a été rapportée. La synaptogamine peut également lier trois autres protéines que sont la neuorexine, la syntaxine et ap2. Compte tenu de la disparition brutale et précoce de toute homologie entre les séquences identifiées et la famille des synaptogamines, il est possible que la séquence identifiée présentent une délétion par rapport à la séquence naturelle. Pour vérifier cette hypothèse et valider les séquences une expérience de RT-PCR et séquençage a été réalisée en utilisant la séquence de 1644 pb. La séquence obtenir comporte une ORF de 420aa présentant une homologie de même ordre avec les synaptogamines.

Pour tenter d'obtenir une séquence plus grande, et vérifier si la séquence obtenue pouvait correspondre à une forme d'épissage, une expérience d'élongation par 5'-RACE a été initiée à partir de la région 3' de la séquence validée, au moyen des oligos L1 et L2 sur une préparation d'ADNe de poumon humain.

5

Les résultats obtenus sont présentés sur la figure 2 et montrent l'identification de 8 clones correspondant à 6 extrémités 5' terminales différentes. Trois d'entre elles contiennent un codon STOP qui interrompt l'ORF (clones A12, F2, F12) et le clone A3 ne contient aucune ORF. La présence de différents transcrits a été confirmée par

10 RT-PCR et nested RT-PCR (Table 1).

Table 1

RT-PCR	Primaire	PCR Secondaire U3-L3	PCR Secondaire U1-L4	PCR Secondaire C-B
U3-L3	170			
A-L4	153		+	
A-L3	Smear	+		
U1-L4	130			
U1-L3	Smear	+		
U1-B	415			+
U2-B	515			+
Taille attendue		170	130	120

Les couples d'amorces U3-L3 et C-B sont spécifiques du fragment commun de la séquence, les oligo A et U1 sont spécifiques de la séquence initiale et du clone C11, l'oligo L4 est spécifique de la séquence initiale et l'amorce U2 est spécifique du clone A3. Une seconde 5'-RACE a été réalisée avec les oligos L3 et L7 localisés dans la région commune des différents clones (Figure 2). Les résultats obtenus sont présentés dans les Figures 3 et 4. La présence de différents transcrits a été confirmée par RT-PCR et nested RT-PCR (Table 2).

20

Table 2

RT-PCR	Résultat	PCR Secondaire C-B	PCR Secondaire U3-B	PCR Secondaire U5-L7
U4-F	Smear	+	+	+
U5-F	Smear	+	+	+
U3-F	1550 pb	+	+	
Taille attendue (pb)		120	385	530

La séquence des amorces et oligonucléotides est donnée dans les Tableaux 3 et 4

5 (SEQ ID NO : 16-37).

Tableau 3

				SEQ ID
LY111_U4	GCAGTTCTGCCTGTTTCATC	23	à 41	16
LY111_U5	TTCAAAACACAGAGGAGGAG	319	à 338	17
LY111_U3	GAATTTGGTCAGTTTAGAGG	759	à 778	18
LY111_L7	TTCTGGGATTTGGAGAGCTTTTTCAC	851	à 825	19
LY111_L6	TCTGTCTGTCCACAGACTGCC	914	à 892	20
LY111_L3	GACTGGCTCCGTCTCTCTG	928	à 910	21
LY111_C	AAGCAACAGAATCTCCCATCC	1029	à 1049	22
LY111_B	GCATTGTCAAAATTGCCATC	1147	à 1127	23
LY111_E	AGGCGGAGAAATACGAAGAC	1543	à 1562	24
LY111_D	GCAGAGTGAGACAGCCCTTAAC	1767	à 1746	25
LY111_L2	CTTCCTCAGGACTGGCGACTTCAG	1811	à 1782	26
LY111_L1	CAAGCGGTCTGTTTCATTCCAAAGAG	1934	à 1913	27
LY111_F	AAGAGGAGATAACCCACCAGAG	2288	à 2269	28

Tableau 4

LY111_A	TCGTAGAGCAGCAGGTCCAAG	14	à 34	46
LY111_U1	AGGGCTGCTGGCTATTTTTTC	36	à 55	29
LY111_L4	TAAGAAATGGGTTGTGAAC	148	à 166	30
LY111_C	AAGCAACAGAATCTCCCATCC	1029	à 1049	31
LY111_B	GCATTGTCAAAATTGCCATC	1147	à 1127	32
LY111_E	AGGCGGAGAAATACGAAGAC	1543	à 1562	33
LY111_D	GCAGAGTGAGACAGCCCTTAAC	1767	à 1746	34
LY111_L2	CTTCCTCAGGACTGGCGACTTCAG	1811	à 1782	35
LY111_L1	CAAGCGGTCTGTTTCATTCCAAAGAG	1934	à 1913	36
LY111_F	AAGAGGAGATAACCCACCAGAG	2288	à 2269	37



L'ensemble de ces résultats permet de valider la séquence consensus correspondant à l'isoforme longue (Figure 9, SEQ ID NO :12 et 13) et à l'isoforme courte (Figure 10, SEQ ID NO :14 et 15) de la protéine PAPI identifiée à partir de poumon humain. Cette protéine est également désignée dans la suite des exemples par le terme LY111.

- 5 L'isoforme long est codé par une ORF de 1833 pb localisée aux résidus 237-2069 de SEQ ID NO :12, et comprend 610 acides aminés. Le signal de polyadénylation est localisé à partir du nucléotide 2315. L'isoforme court est codé par une ORF de 942 pb localisée aux résidus 429-1370 de SEQ ID NO :14, et comprend 313 acides aminés. Le signal de polyadénylation est localisé à partir du nucléotide 1616.

10

Des expériences de Northern blot ont ensuite été réalisées sur différents tissus humains avec des sondes (amplifier CD et E-F) et ont permis de révéler un transcrit de 6 kb dans le muscle, un transcrit dans le cœur (3kb), ainsi qu'un transcrit de 6 kb dans le foie fœtal. L'Exemple 7 décrit par ailleurs le clonage d'un transcrit dans le

15

cerveau fœtal humain. Différentes études d'homologies ont été effectuées dans différentes bases de données protéiques, dont les résultats sont présentés dans le tableau 5 suivant.

Tableau 5

Banque	Homologie
Genpept1116	G5926736 (AB025258) granuphilin-a Identité : 31% (215/679), Homologie (POS) : 46% (322/679)
	G5926738 (AB025259) granuphilin-b Identité : 31% (150/479), Homologie (POS) : 47% (230/479)
	G1235722 (D70830) Doc2 beta (homo sapiens) Identité : 25% (74/292), Homologie (POS) : 43% (127/292)
	G289718 (L15302) Synaptogamin-I Identité : 26% (77/293), Homologie (POS) : 45% (133/293)
Swissprot	SP :SYT1_CAEEL Synaptogamin I Identité : 26% (77/293), Homologie (POS) : 45% (133/293)
	SP :SYT2_MOUSE Synaptogamin II Identité : 24% (72/293), Homologie (POS) : 44% (131/293)

20

**EXEMPLE 7 : CLONAGE DE DEUX TRANSCRITS *FULL-LENGTH* DE PAPI (LY111B) A PARTIR D'ADN COMPLEMENTAIRE DE CERVEAU FETAL HUMAIN**

- 5 Afin de confirmer la présence d'un transcrit Ly111b *full-length* (« complet ») dans le cerveau humain, une PCR a été effectuée à partir d'ADN complémentaire issu de cerveau fœtal humain (Marathon Ready cDNA, Clontech), en utilisant comme amorces les oligonucléotides LyF1 (AAT GGA AGG GCG TGA CGC, figure 5, SEQ ID NO : 38) et HA71 (CCT CAC GCC TGC TGC AAC CTG, SEQ ID NO : 39). Un
- 10 fragment d'ADN faiblement représenté d'environ deux kilobases a été amplifié. Le produit de cette première PCR a servi de matrice pour une PCR *nested*, effectuée avec les oligonucléotides LyEcoF (GCACGAATTC ATG GCC CAA GAA ATA GAT CTG, SEQ ID NO : 40) et HA72 (CTG TCT TCG TAT TTC GCC TTG, SEQ ID NO : 41). Les produits amplifiés ont été digérés avec les enzymes de restriction
- 15 EcoRI (intégré dans l'oligonucléotide LyEcoF) et BstEII (figure 5) et insérés dans le vecteur d'expression pcDNA3, puis leur séquence a été déterminée. L'analyse de la séquence des clones obtenus a révélé la présence de deux transcrits Ly111b potentiels *full-length* dans le cerveau fœtal humain (figure 5). Le premier de ces transcrits (Ly111b<sub>fullA</sub>) correspond à l'ARNm identifié dans le poumon humain (Exemple 6) et
- 20 code une protéine de 609 acides aminés (pLy111b<sub>fullA</sub>; figures 5, 6, SEQ ID NO : 42-43). Le deuxième (Ly111b<sub>fullB</sub>) représente vraisemblablement un produit d'épissage alternatif d'un ARNm primaire commun. Dans ce transcrit, identique à Ly111b<sub>fullA</sub>, la séquence entre les nucléotides 752 et 956 de la séquence validée dans le poumon humain est absente (SEQ ID NO : 42). Ly111b<sub>fullB</sub> code ainsi une protéine de 541
- 25 acides aminés (pLy111b<sub>fullB</sub>) identique à pLy111b<sub>fullA</sub>, dans laquelle cependant le domaine inclus entre les acides aminés 172 et 240 (figures 5, 7, SEQ ID No : 44-45) vient à manquer. Les deux protéines pLy111b<sub>fullA</sub>-pLy111b<sub>fullB</sub> intègrent le domaine d'interaction avec le fragment de la Parkine comprenant les acides aminés 135 à 290, identifié chez la levure (séquence initiale Ly111b, figure 5), et peuvent donc
- 30 théoriquement maintenir cette interaction.

Les protéines pLy11b<sub>fullA/fullB</sub> appartiennent à la famille RIM/Rabphilinc

pLy11b<sub>fullA/fullB</sub> présentent une homologie avec les protéines de la famille RIM/rabphilinc (Wang Y, Sugita S & Sudhof TG. The RIM/NIM family of neuronal C2 domain proteins. *J Biol Chem* (2000) 275, 20033-20044) et en particulier avec les granulophilines (Wang Jie, Takacuchi T, Yokota H & Izumi T. Novel Rabphilin-3-like protein associates with insulin-containing granules in pancreatic beta cells. *J Biol Chem* (1999) 274, 28542-28548). Elles se caractérisent par la présence d'un domaine *zinc finger* dans la partie N-terminale et de deux domaines C<sub>2</sub> dans la partie C terminale (Figures 6 et 7). Le domaine *zinc finger* des protéines de la famille RIM/rabphilinc a été impliqué dans l'interaction avec les protéines Rab. Ces dernières, des protéines qui fixent le GTP, sont des composantes essentielles de la machinerie du trafic membranaire dans les cellules eucaryotes. D'autre part, il a été décrit que les domaines C<sub>2</sub> des protéines de la famille RIM/rabphilinc peuvent lier des membranes à travers l'interaction avec des phospholipides.

Expression des protéines pLy11b<sub>fullA/fullB</sub> dans les cellules de la lignée cos-7 : co-localisation avec la Parkine

La séquence codante des transcrits Ly11b<sub>fullA/B</sub> a été insérée dans le vecteur d'expression eucaryote pcDNA3 en phase avec la séquence codant un épitope myc N-terminal (pcDNA3-mycLy11b<sub>fullA/B</sub>). Des cellules de la lignée cos-7 transfectées à l'aide de ces vecteurs produisent des protéines d'un poids moléculaire apparent d'environ 67 kDa (pcDNA3-mycLy11b<sub>fullA</sub>) et 60 kDa (pcDNA3-mycLy11b<sub>fullB</sub>), correspondant au poids moléculaire attendu. Ces protéines, mises en évidence par immunomarquage à l'aide d'un anticorps dirigé contre l'épitope N-terminal myc, se distribuent de manière non homogène, ponctuelle, dans le cytoplasme, les prolongements et parfois le noyau des cellules de la lignée cos-7 (figure 8a, b, colonne A). Lorsqu'elles sont sur-exprimées avec la Parkine, révélée à l'aide de l'anticorps

4

## REFERENCES BIBLIOGRAPHIQUES

- Abbas, N. *et al.* (1999). A wide variety of mutations in the parkin gene are responsible for autosomal recessive parkinsonism in Europe. *Hum Mol Genet.* **8**, 567-574.
- Bartel, P.L. *et al.* (1993). D.A Hartley Ed, Oxford University press, 153
- Chung, CT. *et al.* (1989). One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. *Proc. Natl. Acad. Sci. USA.* **86**, 2172-2175.
- Fields, S. and Song, O. (1989). A novel genetic system to detect protein-protein interactions. *Nature.* **340**, 245-246.
- Gietz, RD. *et al.* (1992). Improved method for high efficiency transformation of intact yeast cells. *Nucleic Acids Res.*, **20**, 1425.
- Hattori, N. *et al.* (1998). Molecular genetic analysis of a novel Parkin gene in Japanese families with autosomal recessive juvenile parkinsonism: evidence for variable homozygous deletions in the Parkin gene in affected individuals. *Ann Neurol.* **6**, 935-941.
- Kahn, A. *et al.* (1991) Thérapie génique : espoirs et limites. *Médecine et Sciences.* **7**, 705-714.
- Kitada, T. *et al.* (1998). Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. *Nature.* **392**, 605-608.
- Leroy, E. *et al.* (1998). The ubiquitin pathway in Parkinson's disease. *Nature.* **395**, 451-452.
- Lewy, F.H. (1912). in *Handbuch der Neurologie* (Lewandowski, M., ed) pp 920 - 933. Springer, Berlin
- Lucking, C. *et al.* (1998). Homozygous deletions in parkin gene in European and North African families with autosomal recessive juvenile parkinsonism. *Lancet.* **352**, 1355-1356.
- Maniatis, T. *et al.* (1989). Molecular cloning, second edition. Cold Spring Harbor Laboratory. Cold Spring Harbor, N.Y.

- Morett, E.** (1999). A novel transactivation domain in parkin. *Trends Biochem Sci.* **24**, 229-231.
- Peytavi, R. et al.** (1999). HEED, the product of the human homolog of the murine *ced* gene, binds to the matrix protein of HIV-1. *J. Biol. Chem.* **274**, 1635-1645.
- 5 **Polymeropoulos, M.H. et al.** (1997). Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. *Science.* **276**, 2045-2047.
- Sanger, F. et al.** (1977). DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* **74**, 5463-5467.
- 10 **Shimura, H. et al.** (1999). Immunohistochemical and subcellular localization of Parkin protein: absence of protein in autosomal recessive juvenile parkinsonism patients. *Ann Neurol.* **45**, 668-672.
- Sunada, Y. et al.** (1998). Differential expression of the parkin gene in the human brain and peripheral leukocytes. *Neurosci Lett.* **254**, 180-182.
- 15 **Vojtek, AB. et al.** (1993). Mammalian Ras interacts directly with the serine/threonine kinase Raf. *Cell.* **74**, 205-214.

# REVENDECATIONS

1. Composé capable de moduler, au moins partiellement, l'interaction entre la protéine PAP1 ou un homologue de celle-ci, et la Parkine.

5 2. Composé selon la revendication 1, caractérisé en ce qu'il ralentit, inhibe ou stimule, au moins partiellement, ladite interaction.

3. Composé selon l'une des revendications 1 ou 2, caractérisé en ce qu'il est capable de lier le domaine d'interaction entre la protéine PAP1 ou un homologue de celle-ci, et la Parkine.

10 4. Composé selon l'une des revendications 1 à 3, caractérisé en ce qu'il s'agit d'un composé de type peptidique, nucléique, lipidique, saccharidique ou d'un anticorps.

5 5. Composé selon la revendication 4, caractérisé en ce qu'il s'agit d'un composé peptidique comprenant tout ou partie de la séquence peptidique SEQ ID N°2 ou un de ses dérivés.

15 6. Composé selon la revendication 4, caractérisé en ce qu'il s'agit d'un composé peptidique comprenant une région dont la séquence correspond à tout ou une partie fonctionnelle du site d'interaction de la protéine PAP1 avec la Parkine.

20 7. Composé selon la revendication 4, caractérisé en ce qu'il s'agit d'un composé peptidique dérivé de la protéine PAP1 (et/ou des formes homologues) portant une région effectrice rendue non fonctionnelle.

8. Polypeptide comprenant la séquence SEQ ID NO :2 ou un dérivé ou fragment de celle-ci.

25 9. Polypeptide selon la revendication 8, comprenant au moins 5 résidus consécutifs de la séquence SEQ ID NO : 2, de préférence au moins 9, plus préférentiellement au moins 15.

10 10. Polypeptide selon la revendication 8, comprenant tout ou partie de la séquence SEQ ID NO :13, 15, 43 ou 45 ou d'un variant de celles-ci, en particulier au moins 5 résidus consécutifs, de préférence au moins 9, plus préférentiellement au moins 15 résidus consécutifs de la séquence SEQ ID NO :13, 15, 43 ou 45.

11. Acide nucléique codant pour un composé peptidique selon l'une des revendications 4 à 10.

12. Acide nucléique selon la revendication 11, caractérisé en ce qu'il comprend tout ou partie de la séquence SEQ ID NO : 1, 12, 14, 42 ou 44, ou d'une  
5 séquence dérivée de celles-ci.

13. Acide nucléique codant pour un polypeptide selon la revendication 8 ou 11.

14. Acide nucléique, notamment une sonde nucléotidique, capable d'hybrider avec un acide nucléique selon l'une des revendications 11 à 13 ou avec leur brin  
10 complémentaire.

15. Vecteur comprenant un acide nucléique selon l'une des revendications 11 à 14.

16. Virus recombinant déficient comprenant un acide nucléique selon l'une des revendications 11 à 14.

17. Acide nucléique choisi parmi les acides nucléiques de séquence SEQ ID NO : 16-41, 46.

18. Anticorps ou fragment ou dérivé d'anticorps, caractérisé en ce qu'il est dirigé contre un composé peptidique selon l'une des revendications 4 à 10.

19. Anticorps selon la revendication 18, caractérisé en ce qu'il reconnaît un  
20 polypeptide selon la revendication 9 ou 10.

20. Composition pharmaceutique comprenant au moins un composé selon l'une des revendications 1 à 10 ou un anticorps selon la revendication 18 ou 19.



21. Composé non peptidique ou un composé de nature non exclusivement peptidique susceptible de moduler, au moins partiellement, l'interaction de la protéine PAP1 ou un homologue de celle-ci, avec la Parkine.

22. Composé selon la revendication 21 caractérisé en ce que les motifs actifs d'un peptide selon l'une des revendications 5 à 7 ont été dupliqués avec une structure non peptidique ou de nature non exclusivement peptidique.

23. Composition pharmaceutique comprenant au moins un acide nucléique selon l'une des revendications 11 à 14 ou un vecteur selon la revendication 15 ou 16.

24. Composition pharmaceutique comprenant un composé peptidique selon l'une des revendications 4 à 10.

25. Composition selon la revendication 22, 23 ou 24, destinée au traitement de pathologies neurodégénératives.

26. Procédé pour le criblage ou la caractérisation de molécules actives, comprenant une étape de sélection de molécules capables de lier la séquence SEQ ID NO2 ou la séquence SEQ ID NO4, ou un fragment de celles-ci.

27. Procédé pour le criblage ou la caractérisation de molécules actives, comprenant une étape de sélection de molécules capables de lier une séquence choisie parmi SEQ ID NO 13, 15, 43 et 45 ou un fragment de celles-ci.

28. Procédé de production d'un composé peptidique selon l'une des revendications 4, à 10, comprenant la culture d'une cellule contenant un acide nucléique selon l'une des revendications 11 à 14 ou un vecteur selon la revendication 15 ou 16, dans des conditions d'expression dudit acide nucléique, et la récupération du composé peptidique produit.

29. Protéine PAP1 humaine sous forme isolée.

5

31. Mammifère non-humain comprenant dans ses cellules un acide nucléique selon l'une des revendications 11 à 14.

## ABREGE

La présente invention concerne de nouveaux composés et leurs utilisations, notamment pharmaceutiques, diagnostiques ou comme cibles pharmacologiques. Plus particulièrement, la présente invention concerne une nouvelle protéine, désignée PAP], ainsi que de nouveaux peptides et composés capables de moduler au moins partiellement l'activité de la Parkine.